Common species shape the world around us, and changes in their commonness signify large-scale shifts in ecosystem structure and function. However, our understanding of long-term ecosystem response to environmental forcing in the deep past is centred on species richness, neglecting the disproportional impact of common species. Here, we use common and widespread species of planktonic foraminifera in deep-sea sediments to track changes in observed global occupancy (proportion of sampled sites at which a species is present and observed) through the turbulent climatic history of the last 65 Myr. Our approach is sensitive to relative changes in global abundance of the species set and robust to factors that bias richness estimators. Using three independent methods for detecting causality, we show that the observed global occupancy of planktonic foraminifera has been dynamically coupled to past oceanographic changes captured in deep-ocean temperature reconstructions. The causal inference does not imply a direct mechanism, but is consistent with an indirect, time-delayed causal linkage. Given the strong quantitative evidence that a dynamical coupling exists, we hypothesize that mixotrophy (symbiont hosting) may be an ecological factor linking the global abundance of planktonic foraminifera to long-term climate changes via the relative extent of oligotrophic oceans.

1. Introduction

True species richness can be elusive even in well-studied ecosystems, because most species are very rare, and relatively few species account for most of the total abundance [1,2]. For example, only approximately 1.4% of the estimated tree species account for half of the biomass of the Amazon forest and control the cycling of water, carbon and nutrients [3]. Similarly, a recent survey of eukaryotic plankton diversity in the global ocean found that approximately 0.24% of the taxa accounted for half of the total number of rDNA reads [4]. The relative richness of major trophic groups in the plankton is similar across ocean basins, and codiversification of protistan parasites, mutualistic symbionts and their hosts suggests an important role for ecological (‘biotic’) interactions in driving global plankton diversity [4,5]. Oceanographic (‘abiotic’) factors, on the other hand, such as differences in nutrient levels among ocean basins, are more clearly reflected in the abundance of major groups [4,5]. For example, the species richness of diatoms is comparable across nutrient-rich coastal waters and oligotrophic oceans, but their abundance is strikingly different in high- and low-nutrient settings [6]. These first-order patterns of richness and abundance in global eukaryotic plankton suggest that long-term environmental
forcing of global abundance, such as an increase in the global abundance of low-nutrient adapted groups in response to an expansion of oligotrophic conditions, can be independent of any change in richness.

Major oceanographic regime changes, linked to climatic and tectonic processes, have occurred on geological time scales. Geological archives may thus provide insights into how global environmental changes have impacted ocean biomes, as preserved in the sedimentary record. Declining abundance of common species may have dramatic impacts on ecosystem functioning [7–9], with potential cascading effects onto biogeochemical cycling and biotic mitigation of environmental perturbations. However, efforts to link past climatic and oceanographic changes to planktonic biota in deep time have remained focused on temporal changes in species richness [10–12], and the importance of common species as tracers of past ecosystem response has been neglected.

We seek to redress the situation by targeting species that are globally common and widespread at any given time, using the Summed Common species Occurrence Rate (SCOR). SCOR sums the observed occupancy (the proportion of sampled sites at which a species is present and observed) across a species set and is driven by the most widespread species in the set [13]. In general, widespread species are also abundant [14], and a positive relationship between global abundance and occupancy is found in the global ocean plankton [4], including the planktonic foraminifera (electronic supplementary material, figure S10). Here, we define commonness as high global site occupancy, and show that if the global abundance–occupancy relationship is positive (albeit noisy), then SCOR is a proxy for relative changes in global abundance integrated over the total set of species included. We do not consider local abundances or geographical ranges of individual species. A species contributes to SCOR only if and when it is widespread, which accommodates the geological transience of commonness; in the fossil record, species and higher taxa generally show a gradual expansion and contraction of their spatial extent of occurrence, being rare (restricted) in the early and late stages of their known stratigraphic range [15–17]. Furthermore, species that make up the bulk of the fossil record, and biostratigraphically useful species in particular, tend to be common and widespread.

Turning to the rich deep-sea sedimentary record of the Cenozoic era (0–65 Ma), we apply SCOR to global occurrences of planktonic foraminifera, a calcifying unicellular zooplankton group of central importance to our understanding of Cenozoic climate history. The taxonomy of planktonic foraminifera is based on test (shell) morphology, and despite the presence of cryptic diversity, the validity of extant morphological taxa is generally supported by ribosomal sequence data [18,19]. Global patterns of morphospecies richness of extant planktonic foraminifera have been linked to various environmental factors, including sea surface temperature, mixed layer depth, productivity, geography and ocean currents [20,21]. On Cenozoic time scales, fossil diversity dynamics in planktonic foraminifera have been linked to a combination of ecological traits and oxygen isotope records via statistical model selection [10], or correlated to changes in deep-sea sedimentation [22]. The dynamics of global commonness, on the other hand, have remained unexplored.

We propose that changes in the commonness of the most common species may represent a more tractable and meaningful measure of past ecosystem response to environmental forcing than those in the estimated total number of species. As a measure of potential physical forcing, we use a reconstruction of changes in deep-ocean temperature (DOT) throughout the Cenozoic [23]. We use the DOT record because it is a robust, globally meaningful signal of long-term changes, compatible with our integrated, global-scale analysis. Temperature proxy records for the surface ocean show considerable spatial heterogeneity [24] and would call for a spatially explicit analysis, which is not considered here. Geological proxy records are generally noisy mixtures of signals that represent multiple processes and are derived from a sedimentary record that is itself an active component of the Earth system. Any causal connection detected between proxy records would necessarily be indirect with respect to the underlying processes of interest. Nonetheless, the DOT proxy record of high-latitude and deep-water temperature reflects a set of interlinked, climate-related oceanographic variables, including changes in geographical and vertical thermal structure, considered to be important physical controls on the long-term evolution of planktonic foraminifera [10,22,25–27]. Here, we find evidence of a dynamical coupling between planktonic foraminifera SCOR and DOT using three conceptually very different methods for detecting causality in time series.

2. Material and methods

(a) Data

Occurrences of macro- and microperforate planktonic foraminifera were retrieved from the Neptune Sandbox Berlin (NSB) database [28,29] (on 22 April 2015). SCOR was calculated using 1.5 Myr time bins, in order to maximize the number of temporal observations for time-series analysis, but our results also hold on a lower resolution of 0.5 Myr. In addition to a SCOR curve for 572 species in the original NSB taxonomy (updated in 2014 to match the taxonomic name list of the International Ocean Discovery Program), we calculated an ensemble of 20 SCOR curves that included from 51 to 100 of the most common species. To assess the robustness of SCOR to differences in taxonomic opinion, we established the set of 100 most common species using a different taxonomic scheme [30] and checked that the time intervals of commonness for these species in NSB were congruent with published stratigraphic ranges of each species after this taxonomic re-mapping [30–32] (electronic supplementary material, figure S2 and table S1).

DOT estimates were obtained from Cramer et al. [23], using the 18O record (based on subtracting New Jersey sea-level estimates from a benthic δ18O stack) for the interval 9–65 Ma, and the scaled 818O record for the interval 0–9 Ma, with their interpolation at 0.1 Myr resolution. The SCOR and DOT data are available in the electronic supplementary material.

(b) Summed common species occurrence rate

We treat the observation of a specific number of individuals as a Poisson-distributed variable with parameter λ in each time bin. The probability of finding an individual of species i in time bin j is then \( p_{ij} = 1 - \exp(-\lambda_{ij}) \), and thus \( \lambda_{ij} = -\ln(1 - p_{ij}) \). In practice, \( p_{ij} \) is estimated as \( y_{ij}/n_{ij} \), where \( y_{ij} \) is the number of sites in which species i is observed at time bin j and \( n_{ij} \) is the number of sites in that time bin where at least one of the species included in the analysis is observed. SCOR is the total density of a given set of \( m_{ij} \) species in a time bin,

\[
\text{SCOR}_i = \sum_{j=1}^{m_{ij}} \lambda_{ij}.
\]
and we estimate its variance by the delta method [33],

$$\text{Var}(\text{SCOR}_j) = \sum_{i=1}^{n} \frac{p_i}{(1-p_i)\eta_j}$$ \hspace{1cm} (2.2)

If species $i$ occurs at all sites in time $j$, then $\lambda_{ij}$ is undefined. As $p$ approaches 1, $\lambda$ grows logarithmically, hence widespread species dominate SCOR.

(c) Poseidon simulations

Here, we evaluate the sensitivity of SCOR to changes in (i) spatial sampling completeness and (ii) the shape of the species rank-abundance distribution (RAD). Our simulation model (Poseidon) consists of a user-specified number of spatial cells (sites), number of time steps, total number of species (richness) and total number of individuals (abundance). In our simulations, we allow true richness and abundance to vary independently (figure 1a), to highlight the key point that SCOR is unaffected by changes in richness. In each time step, we distribute individuals randomly among species, obeying a lognormal RAD. The RAD shape parameter ($\alpha$) can be fixed or time-varying. We distribute individuals randomly among sites, and we sample only a proportion of the sites. This proportion increases from 0.1 to 0.4 through time (sampling coverage decreases with age), with any short-term sampling variability superimposed on this trend (figure 1b). Local preservation variability can be subsumed in the spatial sampling, because the quantities we focus on here are based on species occurrence (presence/absence) at sites, not on local abundance of individuals.

To test our proposition that commonness can be more tractable than richness, we compare the performance of SCOR to richness estimators commonly used in palaeobiology. Poseidon therefore calculates raw richness ($S$), range-through richness (RT; assuming a species existed in all time bins between its first and last occurrence), Shannon entropy ($H$), and several sampling-standardized richness metrics, including classic richness. In each time step, we distribute individuals randomly among species, obeying a lognormal RAD. The RAD shape parameter ($\alpha$) can be fixed or time-varying. We distribute individuals randomly among sites, and we sample only a proportion of the sites. This proportion increases from 0.1 to 0.4 through time (sampling coverage decreases with age), with any short-term sampling variability superimposed on this trend (figure 1b). Local preservation variability can be subsumed in the spatial sampling, because the quantities we focus on here are based on species occurrence (presence/absence) at sites, not on local abundance of individuals.

To test our proposition that commonness can be more tractable than richness, we compare the performance of SCOR to richness estimators commonly used in palaeobiology. Poseidon therefore calculates raw richness ($S$), range-through richness (RT; assuming a species existed in all time bins between its first and last occurrence), Shannon entropy ($H$), and several sampling-standardized richness metrics, including classic richness. In each time step, we distribute individuals randomly among species, obeying a lognormal RAD. The RAD shape parameter ($\alpha$) can be fixed or time-varying. We distribute individuals randomly among sites, and we sample only a proportion of the sites. This proportion increases from 0.1 to 0.4 through time (sampling coverage decreases with age), with any short-term sampling variability superimposed on this trend (figure 1b). Local preservation variability can be subsumed in the spatial sampling, because the quantities we focus on here are based on species occurrence (presence/absence) at sites, not on local abundance of individuals.

To test our proposition that commonness can be more tractable than richness, we compare the performance of SCOR to richness estimators commonly used in palaeobiology. Poseidon therefore calculates raw richness ($S$), range-through richness (RT; assuming a species existed in all time bins between its first and last occurrence), Shannon entropy ($H$), and several sampling-standardized richness metrics, including classic richness. In each time step, we distribute individuals randomly among species, obeying a lognormal RAD. The RAD shape parameter ($\alpha$) can be fixed or time-varying. We distribute individuals randomly among sites, and we sample only a proportion of the sites. This proportion increases from 0.1 to 0.4 through time (sampling coverage decreases with age), with any short-term sampling variability superimposed on this trend (figure 1b). Local preservation variability can be subsumed in the spatial sampling, because the quantities we focus on here are based on species occurrence (presence/absence) at sites, not on local abundance of individuals.

To test our proposition that commonness can be more tractable than richness, we compare the performance of SCOR to richness estimators commonly used in palaeobiology. Poseidon therefore calculates raw richness ($S$), range-through richness (RT; assuming a species existed in all time bins between its first and last occurrence), Shannon entropy ($H$), and several sampling-standardized richness metrics, including classic richness. In each time step, we distribute individuals randomly among species, obeying a lognormal RAD. The RAD shape parameter ($\alpha$) can be fixed or time-varying. We distribute individuals randomly among sites, and we sample only a proportion of the sites. This proportion increases from 0.1 to 0.4 through time (sampling coverage decreases with age), with any short-term sampling variability superimposed on this trend (figure 1b). Local preservation variability can be subsumed in the spatial sampling, because the quantities we focus on here are based on species occurrence (presence/absence) at sites, not on local abundance of individuals.
rarefaction (CR), occurrences-squared weighted (O2W) [34] and two versions of subsample quorum subsampling [35,36]; SQS 3.3 (R script downloaded from the SQS website on 26 August 2014) and SQS 4.3 (perl script kindly provided by J. Alroy on 3 September 2014). We ran SQS 3.3 on sampled individuals (abundances) and SQS 4.3 on formatted Poseidon species occurrences, with a quorum level of 0.6 in all runs. Both SQS versions yielded very similar results when given the same type of data (abundances or occurrences). We used 100 iterations/trials in all subsampling methods (CR, O2W and SQS). Shannon H from the SQS 3.3 R script was used to calculate Piolou’s J evenness. Goodness of fit between true and estimated richness, and between true abundance and SCOR, was measured by the coefficient of determination ($R^2$) between time series. A link to the Poseidon R scripts is provided below.

### (d) Time-series analysis

If we take seriously the notion that climate-related oceanographic changes and planktonic ecosystems are interacting components of the Earth system, then we need to approach the possible long-term linkage between SCOR and DOT from a dynamical systems perspective. However, SCOR and DOT are both indirect proxy records that are not uniquely determined by any single physical or ecological process, and a dynamical coupling detectable in the proxy records would be a result of transitive chains of unobserved causal intermediates. In addition, the tectonic, climatic, atmospheric and biotic boundary conditions have all evolved over the Cenozoic. These challenges, combined with uncertainty in the relative importance and scales of mechanisms, limit the feasibility of explicit process modelling of a SCOR–DOT interaction throughout the Cenozoic. Instead, our approach is to test the hypothesis of a causal relationship using dynamical information inherent in the observed records, which enables us to advance beyond static correlations while avoiding unwarranted mechanistic assumptions.

We used three different time-series analysis methods to test for a causal relationship between SCOR and DOT: (i) convergent cross mapping (CCM) [37,38], based on the concept of state space reconstruction from time-delay embedding [39]; (ii) information transfer (IT) analysis [40,41], based on the concept of directional information flow [42]; and (iii) Bayesian inference of causal connections based on linear stochastic differential equations (SDEs) [43,44]. To distinguish unidirectional forcing from bidirectional causality, we calculated CCM for different time lags of the original time series. If there is a discernible delay between cause and effect, then CCM is expected to peak for negative time lags in the direction(s) of true causality (past predicts future). If true causality is unidirectional but with synchrony (inducing two-way predictability), then any CCM skill in the non-causal direction is expected to peak for positive lags (future predicts past). IT analysis has previously been applied to geological records [13,41,45]. Here, we expand on earlier applications by using time lags analogously to the lagged CCM. If there is a causal delay, then IT is expected to peak for negative time lags in the direction(s) of true causality (past — future), and any IT in the non-causal direction is expected to peak for positive lags. In the analysis of SCOR and DOT, lagged CCM and IT are both reported as median values and 95% ranges at each lag calculated by iterative sampling across an ensemble of SCOR curves, with significance evaluated against a distribution of surrogate time series. SDE analysis is a model-based approach. Here, we first determined the most probable stochastic models for SCOR and DOT separately, including possible hidden (unobserved) process layers, then we determined the most probable connections between these two models, distinguishing correlations from causal connections and inferring the relative strength, direction and response time of causal connections [43]. A more detailed description of each time-series analysis method is provided in the electronic supplementary material.

### 3. Results

#### (a) Poseidon numerical experiments

We find that SCOR is robust to variability in both spatial sampling and RAD shape (figure 1d), as well as variability in the abundance–occupancy relationship (electronic supplementary material, figure S1). As expected, raw S decays rapidly with increasing sampling variability, but is insensitive to changes in RAD shape ($\sigma$). RT is relatively robust to both factors, as the species pool and/or sampling is sufficient to avoid severe edge effects. Shannon H reflects both richness and evenness, and, with increasing $\sigma$ variability, H becomes more sensitive to $\sigma$ than to changes in richness. O2W shows overall poor agreement with true richness. CR and SQS, being sampling standardization methods, are robust to the isolated effect of spatial sampling variability on richness, all else being equal. However, both CR and SQS are highly sensitive to changes in RAD shape. As with Shannon H, increasing $\sigma$ variability causes CR and SQS to lose track of richness (figure 1d) and to respond to changes in $\sigma$ instead (electronic supplementary material, figure S3). In Poseidon, $\sigma$ variability is a white noise, entirely random with respect to both richness and abundance. Hence, the sensitivity of subsampling methods to RAD shape exists independently of how one might choose to measure sample evenness.

Estimates of SQS richness in the NSB data confirm the sensitivity to RAD shape found in Poseidon. Indeed, SQS richness can be reproduced by simply combining the raw richness and evenness curves, a relationship that holds for different taxa in NSB and across all Poseidon simulations (electronic supplementary material, figure S4). Changes in RAD shape are thus a significant confounding factor, at least for the specific subsampled richness algorithms tested here.

A data source like NSB, in which species occurrences were originally recorded for biostratigraphic purposes, is well suited for SCOR, because SCOR is only sensitive to common and widespread species, and biostratigraphically useful species are typically common and widespread. Nevertheless, we have included a Poseidon experiment where the recording of species (presence/absence) is random with respect to global commonness (electronic supplementary material, figure S5). Even in this scenario, which is designed to be particularly biased against it, SCOR still performs better than richness estimators. This experiment shows that we do not have to capture every abundant species for our approach to be meaningful.

#### (b) Planktonic foraminifera summed common species occurrence rate

Planktonic foraminifera SCOR converges rapidly as species are added in rank order of decreasing commonness (figure 2a). The SCOR curve for the 572 original NSB species is statistically indistinguishable from the SCOR curve of the approximately 100 most common species, because SCOR is driven by the most common and widespread species. We based our final analyses on an ensemble of 20 SCOR curves, including from 81 to 100 of the most common species, after taxonomic remapping (electronic supplementary material, figure S2 and table S1). This SCOR ensemble was then compared with the Cenozoic DOT reconstruction (figure 2b) and subjected to a series of analyses designed to formally detect dynamical causal coupling.
Although the net trends are inversely related, fluctuations in SCOR and DOT do not show consistent negative or positive covariation throughout the Cenozoic (figure 2; for first-differenced data at zero lag, Spearman’s $r = 0.11$, $p = 0.13$). However, CCM skill from foraminifera SCOR to DOT peaks at a negative lag, indicating that the SCOR signal carries a response to past changes captured in the DOT record (‘SCOR $\times$map DOT’; figure 3a). CCM skill in the opposite direction peaks at positive lags, but this skill is not significant relative to the surrogates (figure 3a). This result is consistent with a unidirectional forcing of SCOR by DOT, where predictability potentially goes both ways, but the causal direction peaks at negative lags [38]. IT analysis supports this causal directionality: predictive information flow is significant from past DOT to SCOR, peaking at a negative lag of two to four time bins (approx. 1 Myr; figure 3b). In the opposite direction, IT peaks at the corresponding positive lags, also consistent with unidirectional forcing. Both IT and CCM indicate a time-delayed coupling, but age model noise and the indirect nature of proxy records imply that the absolute magnitude of the lags cannot be given a strict physical interpretation. We therefore limit our interpretation to the sign of the peak lag to infer the directionality of the causal relationship.

Using a series of linear SDEs to model correlation and causality between the two records, we recover relatively strong evidence that SCOR responds to changes in DOT, with a time lag of 0.33–1.1 Myr, in agreement with the CCM and IT analyses. Linear SDE analyses on SCOR and DOT individually show that the best descriptions for SCOR and DOT are a single-layer and a three-layer model, respectively (electronic supplementary material). We allowed for the possibility of a two-way connection in a statistical sense,
because both proxies derive from the deep-sea carbonate record, which is a responsive component of the climate system and could, in principle, impart a shared signal. Consequently, there are 15 possible connection models, including the null hypothesis of no relationship (electronic supplementary material, figure S6). Posterior model probabilities provide substantial evidence for a causal connection between SCOR and DOT (electronic supplementary material). In the two most probable models (figure 3c), although a two-way connection is statistically established, the influence of DOT on SCOR is overwhelmingly dominant (electronic supplementary material). The third most likely model is a simple one-way connection from DOT to SCOR (figure 3c), consistent with the overall causal inference.

4. Discussion

Occurrences in the NSB database stem from ocean drilling expedition shipboard data, and the presence of taxonomic and stratigraphic errors can be problematic for constructing range-based richness or phylogeny [30]. Our approach is much less sensitive to taxonomic and stratigraphic noise.

Furthermore, we do not use local abundance data in NSB, which may be influenced by local fluxes and preservation, to reconstruct global abundance. Instead, SCOR is based only on global occupancy, integrated across widespread species, because globally widespread species are typically globally abundant (electronic supplementary material, figure S1a) [4]. As predicted from the hollow shape of species-abundance distributions, our results are reproducible with fewer than 100 of the most common species (figure 2a).

The overwhelming majority of species in NSB thus do not contribute to SCOR, and the use of different taxonomic schemes does not affect our results. SCOR does not depend on knowing true times of speciation and extinction; a species only contributes to SCOR if/when it is widespread, and stratigraphic outliers or long tails of occurrences, erroneous or not, have no impact. Hence, we need not distinguish between a species having gone truly extinct and it having become a rare, ecological ‘ghost’, like the coccolithophore Tergestiella/Cyclagelosphaera, recently found living 54 Myr after its presumed extinction [46]. In addition, even if morphospecies do not conform strictly to biological species, they may represent successful phenotypes and/or traits that potentially...
link global abundance to ecosystem functioning. We use SCOR for its simplicity and consider it a starting point for more sophisticated occupancy-based approaches [47].

The planktonic foraminifera SCOR does not show positive covariation with a Cenozoic SCOR curve for the coccolithophores, another important calcifying plankton group. Instead, their raw patterns are negatively correlated (Spearman’s $\rho = -0.28$) and first differences are not significantly correlated ($\rho = 0.15$, $p > 0.06$). Although preservation can affect the local relative abundance of these two groups in the sediment [48], it does not explain the different SCOR patterns, because expansion or contraction of the global extent of deep-sea carbonate deposition would cause the preserved global occupancy of both calcifying groups to increase or decrease in similar ways. Aspects of the Cenozoic coccolithophore SCOR pattern have been linked to proxy records of atmospheric CO$_2$ [13] and suggest that these calcifying phytoplankton could thrive in a high-CO$_2$ world [49]. Cooling and CO$_2$ decline across the Eocene–Oligocene transition were accompanied by a lowering of the carbonate compensation depth, which has been attributed to changes in the supply of weathering products to the ocean [50]. Like the coccolithophores [13], however, planktonic foraminifera show declining SCOR across this transition, which is opposite to that expected if SCORs were biased upwards by enhanced deep-sea preservation. On the other hand, if species absence/presence were random with respect to global abundance, then SCOR would suffer from increased short-term volatility, but Poseidon experiments show that even under such conditions, SCOR can track relative changes in global abundance (electronic supplementary material, figure S5). For the 25 sites common to both databases, the age models for the NSB data may deviate from those in the DOT record, but the age models are not systematically offset (electronic supplementary material, figure S7); hence, the presence of age model noise should weaken any statistical associations and bias against our results. Finally, we note that although the number and global coverage of sites in the NSB data decrease with age, this temporal decay of spatial sampling coverage in itself does not create a systematic bias in SCOR (electronic supplementary material, figure S8).

The congruence of our three independent analyses provides strong quantitative evidence of a dynamical coupling between the observed global occupancy of planktonic foraminifera and past climatic and oceanographic changes captured in the DOT proxy record. A Cenozoic increase in ocean thermal stratification due to polar cooling has long been proposed as a ‘universal driver’ of plankton evolution [27], or even a ‘secular abiotic forcing of plankton evolution’ [49]. Assimilating a Cenozoic DOT proxy in models of planktonic foraminifera diversification has also been found to increase model support [52], on time scales comparable with the predictive lags found here. However, we caution against reducing the DOT–SCOR coupling to any single mechanism operating through the evolving biotic, tectonic and atmospheric boundary conditions of the Cenozoic. Rather than attributing process to temporal correlations or net trends, we use information inherent in the data to demonstrate a dynamical coupling, while avoiding mechanistic assumptions. The magnitude of the predictive lag cannot be interpreted as a direct, physical process–response delay, but accommodates an indirect, heterogeneous causal pathway between DOT and SCOR. Our results instead open up new questions on the long-term dynamics of commonness and its linkage to environmental change. As an example, we briefly consider the ecological characteristics of the species responsible for the SCOR pattern.

Planktonic foraminifera inferred to host symbionts in the open-ocean mixed layer (Ecogroup 1 in [30]) are among the most widespread species that contribute to high SCOR during intervals of relative DOT warmth (i.e. positive deviations from the net DOT trend), especially the early Eocene and the late Miocene–Pliocene (figure 2a; electronic supplementary material, figure S2). By contrast, deeper-dwelling species not known to host symbionts contribute most to SCOR during intervals of relative DOT cooling, such as the latest Eocene–Oligocene and, in the case of thermocline-dwelling species, the mid-Miocene. Symbiont-hosting (mixotrophic) rhizarian protists are surprisingly abundant in oligotrophic oceans today [53,54]. If oligotrophic realms expand with surface warming in the near future [55], then the global abundance of mixotrophs could increase. Greater global abundance of mixotrophs might enhance the efficiency of the biological carbon pump in the ocean, through net increases in body size and faster sinking rates [56]. Although the global biogeochemical impact of planktonic foraminifera may be limited, their outstanding fossil record makes them a prime candidate for studying a planktonic ecosystem response to global environmental changes in the past. Despite the opposite net trends, long-term fluctuations in SCOR and DOT show intervals of positive covariation, suggesting that planktonic foraminifera proliferated in periods of relative DOT warmth. To the extent that long-term patterns in the DOT record reflect Cenozoic mean states of surface water temperature (e.g. early Eocene warming followed by long-term Eocene cooling), and if, on average, warmer water conditions correspond to lower nutrient levels, then mixotrophy may have allowed the global abundance of planktonic foraminifera to increase in a warmer ocean. In the warm early Eocene, for example, oligotrophic open-ocean ecosystems may have reached mid- and high latitudes [57]. If higher global abundance of mixotrophic plankton was associated with greater efficiency of the biological carbon pump [56], then the SCOR pattern would be consistent with the sharpened $\delta^{13}$C depth gradients observed in the warm early–mid–Eocene [58], and the reduction of $\delta^{13}$C depth gradients from the warm Pliocene to the cooler Pleistocene [59]. We propose that mixotrophy has been an ecological factor linking the global abundance of planktonic foraminifera to the relative extent of oligotrophic oceans. Testing this hypothesis calls for spatially resolved proxy data and spatially explicit occupancy modelling, which are out of the scope of this study.

In this paper, we have: (i) highlighted the importance of widespread species as indicators of global ecosystem response; (ii) shown that a simple measure of global occupancy can be used to track changes in global abundance, because widespread species are typically globally abundant (i.e. commonness, as we define it); (iii) shown that in the fossil record, commonness is more tractable than richness, because the same fundamental ecological property that lends power to commonness (i.e. only a few species are abundant and widespread) can be a weak spot for richness estimators; (iv) used dynamical information in global-scale proxy records to rigorously test the coarse-grained hypothesis of a causal relationship between observed global
occupancy of planktonic foraminifera and long-term climate changes captured in DOT records; and (v) based on the quantitative evidence for a causal connection, put forward a more fine-grained hypothesis on the possible role of mixotrophy that we hope deserves future scrutiny.

Data accessibility. The datasets used in this article are included in the electronic supplementary material. The Poseidon code is available on GitHub (https://github.com/bhaniss/poseidon).

Authors’ contributions. B.H. and L.H.L. conceived, designed and coordinated the study. L.H.L. wrote the Poseidon code and scripts for querying the NSB and TARA Oceans databases. K.A.H. and D.D. performed CCM analysis. T.R. performed SDE analysis. B.H. carried out the Poseidon simulations, taxonomic remapping, IT analysis and all additional analyses, and drafted the manuscript. All authors discussed results and interpretations, commented on drafts of the manuscript and gave final approval for publication.

Competing interests. We declare we have no competing interest.

Funding. This work is funded by the Research Council of Norway grant nos. 231259 (B.H.) and 230073 (L.H.L.), and by the Bergen Research Foundation (B.H.).

Acknowledgements. We acknowledge the many contributors to the NSB database. We thank Tom Ezard and Dave Lazarus for comments on earlier versions of this manuscript. Thanks to John Alroy for sharing his SQS 4.3 script and to Raphael Morard for assistance with TARA Oceans data. Open Access was supported by the University of Bergen. Wolfgang Kiessling and two anonymous referees provided helpful comments that improved the manuscript.

References


44. Liew LH, Reitan T, Hamik PG. 2015 Ecological interactions on macroevolutionary time scales: clams and bivalves are more than ships that pass in the night. Ecol. Lett. 18, 1030–1039. (doi:10.1111/ele.12485)


